Hand-waving Refined Algebraic Quantization *

Franz Embacher
Institut für Theoretische Physik
Universität Wien
Boltzmanngasse 5
A-1090 Wien

E-mail: fe@pap.univie.ac.at

UWThPh-1997-22
gr-qc/9708016

Abstract

Some basic ideas of the Refined Algebraic Quantization scheme are outlined at an intuitive level, using a class of simple models with a single wave equation as quantum constraint. In addition, hints are given how the scheme is applied to more sophisticated models, and it is tried to make transparent the general pattern characterizing this method.

*Talk given at the meeting *Modern Modified Theories of Gravitation and Cosmology*, Ben Gurion University, Beer Sheva, June 29–30, 1997.

1 Introduction

In a slight misuse of the purpose of this meeting, I will not talk about a "modified" or "alternative" theory of gravity but rather give a brief introduction into a "modified" or "alternative" quantization method. Although currently being applied with considerable success to the program of quantizing full gravity, this method seems to be unknown to a large fraction of the relativity community. Hence, although I am not a very specialist on this topic, let me please take the opportunity to present the basic ideas of what is called Refined Algebraic Quantization (RAQ). I will put emphasis on some of its aspects that can be understood intuitively and present them in a hand-waving fashion, rather than being mathematically very rigorous. In case I succeed in stimulating some interest, the existing literature will certainly be helpful in getting further.

The RAQ scheme deals with constrained systems. The type of constraints most challenging for gravitational theorists and cosmologists is of course the one appearing in the Hamiltonian formulation of general relativity: quadratic in the momenta. The recent development of the RAQ scheme is to some extent linked with the program of loop quantum gravity[1] (the basic idea dating back to the Sixties [2] and having been re-invented in the Nineties [3, 4, 5]). I will ignore the "loop" issue (and also the issue of the $SU(2)$ connection and the famous spin networks appearing there) since, logically, quantum gravity just provides an area of application of the general set of rules called RAQ. Rather than stating these rules precisely in their most sophisticated form, I will first develop an easy version by treating a class of drastly simplified models. In these models almost all features of the "true" gravity constraints have disappeared, except for one: being quadratic in the momenta. Only afterwards, I will provide some hints towards the exactification of the scheme and its generalization to more realistic systems.

In other words, I consider what may be called minisuperspace models or, mathematically equivalent, the motion of a scalar particle in a space-time background. In such a scenario, one has a fixed background structure, consisting of a finite dimensional manifold $\mathcal{M}$, a Pseudo-Riemannian metric $ds^2 = g_{\alpha\beta}\, dx^\alpha dx^\beta$ (signature $- + \ldots +$) and a real function U. Choosing a coordinate system (x^α) in $\mathcal{M}$, the classical phase space locally consists of pairs (x^α, p_α). There is only one classcial constraint, reading

$$\mathcal{C} \equiv p_\alpha p^\alpha + U = 0\,. \tag{1}$$

Upon substituting $p_\alpha \to -i\partial_\alpha$ and choosing the simplest covariant operator ordering, the classical constraint is promoted into the quantum constraint (wave equation)

$$\widehat{\mathcal{C}}\,\psi \equiv (-\nabla_\alpha \nabla^\alpha + U)\psi = 0\,, \tag{2}$$

the solutions of which we call physical states. (This will be slightly modified and made more precise later on).

In a quantum cosmological context, $\mathcal{M}$ is the minisuperspace manifold and equation (2) is the corresponding Wheeler-DeWitt equation. If $\mathcal{M}$ denotes a space-time manifold, equation (2) is the Klein-Gordon equation for a scalar particle moving in this space-time and feeling a potential U. Hence, one and the same mathematical equation (2) may be interpretated in two ways that are entirely different from each other from the point of view of physics. (In particular, the notion of a class of observers moving on a congruence of world-lines in space-time and being freely adjustable has no counterpart in minisuperspace). I stress this point, because the RAQ scheme may look very strange at first sight when applied to a particle quantization scenario, and it may easier loose this alien appeal when a quantum gravitational (or simpler: a quantum cosmological) context is kept in mind.

So let us start with the wave equation (2). In the traditional Klein-Gordon quantization one is guided by the idea of ψ being the wave function of a particle. The key object to give a suitably defined space of solutions of (2) some additional structure is the indefinite scalar product

$$Q(\psi_1,\psi_2) = -\frac{i}{2}\int_\Sigma d\Sigma^\alpha\,(\psi_1^*\overset{\leftrightarrow}{\nabla}_\alpha\,\psi_2)\,, \tag{3}$$

where Σ is a spacelike hypersurface (with sufficiently regular asymptotic behaviour). The integrand of Q is a conserved current, upon which many of the physical interpretations of ψ are based.

Let us recall that for the case of a massive particle in Minkowski space ($g_{\alpha\beta}=\eta_{\alpha\beta}$ being flat and $U=m^2$ being constant) there is a unique decomposition of any (asymptotically well-behaved) wave function into a positive and a negative frequency part

$$\psi(t,\vec{x}) = \int\frac{d^3k}{(2\pi)^{3/2}\sqrt{\omega(\vec{k})}}\Big(a(\vec{k})e^{i(\vec{k}\vec{x}-\omega(\vec{k})t)}+b(\vec{k})e^{i(\vec{k}\vec{x}+\omega(\vec{k})t)}\Big)\equiv\psi_+(t,\vec{x})+\psi_-(t,\vec{x})\,, \tag{4}$$

where $\omega(\vec{k})=\sqrt{\vec{k}^2+m^2}$. The scalar product (3) then reads

$$Q(\psi_1,\psi_2) = \int d^3k\Big(a_1^*(\vec{k})a_2(\vec{k})-b_1^*(\vec{k})b_2(\vec{k})\Big)\,, \tag{5}$$

the minus sign nicely illustrating the indefiniteness of Q and the way how actually *two* Hilbert space structures (one with positive and one with negative sign) emerge from (2). In the early approaches to the quantum theory of relativistic particles, they would have been called one-particle (and one-antiparticle) Hilbert space, respectively.

On a curved background, an analogous construction becomes ambigous, unless there is a local symmetry with timelike trajectories, thus providing a preferred time coordinate with respect to which the frequency decomposition is defined. (In the

Minkowski space example above, one of course assumes the time coordinate of any intertial frame to do this job. Due to Lorentz invariance, the resulting decomposition (4) is independent of the frame chosen). This ambiguity in quantizing a particle in a generic curved background just means that the definition of what is a particle and what is an antiparticle depends on the observer(s). This is not just a shortcome of the quantization method but reflects physical properties: it is the origin of particle production in curved space-time. However, in lack of a preferred expansion like (5) — which displayed two Hilbert space structures — it is not at all easy to say which of the functions ψ on $\mathcal{M}$ solving the wave equation shall qualify as physical states and which shall not, and usually one has to worry about cumbersome boundary conditions.

The modern way to carry out the particle quantization program in a curved background is not this approach but second quantization. However, quantum cosmology introduces a different physical interpretation of the ingredients of (2), so that one must reconsider the quantization method. Here, the RAQ scheme may be brought into the game as a method which, in minisuperspace models of the type (2), generates a well-defined Hilbert space of physical states in a much more direct way than the Klein-Gordon approach. Also in realistic models, containing more than just one constraint, this scheme seems to be more likely to work than one based on a generalization of (3).

2 Delta function acrobatics

In contrast to the Klein-Gordon quantization based on (3), the RAQ relies on the positive definite inner product

$$\langle\psi_1,\psi_2\rangle = \int_{\mathcal{M}} d^n x\,\sqrt{-g}\,\psi_1^*\psi_2 \tag{6}$$

on the Hilbert space $\mathcal{H}$ of square-integrable functions on $\mathcal{M}$. Although this integral will in general not converge if solutions of (2) are inserted, the main motivation driving one away from the traditional scheme is to reject any indefinite scalar product and to retain (6) as the starting point.

Proceeding heuristically, I write down a solution to the quantum constraint (2) in a very formal way. Let us insert the constraint operator $\widehat{\mathcal{C}}$ from (2) into the delta function and arrive at the object $\delta(\widehat{\mathcal{C}})$. This is not as weird as it might look at first sight, because it is actually defined as

$$\delta(\widehat{\mathcal{C}}) = \int_{-\infty}^{\infty}\frac{d\lambda}{2\pi}\,e^{i\lambda\widehat{\mathcal{C}}}\,, \tag{7}$$

in a sense to be specified in a moment. Whenever this construction makes sense, one expects $\widehat{\mathcal{C}}\,\delta(\widehat{\mathcal{C}})=0$, so that, once a function $\psi\in\mathcal{H}$ is specified, the quantity

$$\psi_{\rm phys} = \delta(\widehat{\mathcal{C}})\,\psi\,, \tag{8}$$

provided it exists, should be a solution of the wave equation. Of course, this has to be made more precise. Equation (7) in fact tells us how to proceed: By simply *first* applying $e^{i\lambda\widehat{\mathcal{C}}}$ to ψ and *thereafter* integrating over λ, the quantity (8) becomes an integral that has a chance to converge. A technically important requirement when doing so is that the constraint operator $\widehat{\mathcal{C}}$ shall be self-adjoint, so that $e^{i\lambda\widehat{\mathcal{C}}}$ is unitary for any real λ. This is in fact the case for a wide class of models $(\mathcal{M}, g_{\alpha\beta}, U)$. Moreover, if the spectrum of $\widehat{\mathcal{C}}$ is purely continuous near zero, the definition (7) should intuitively be as reasonable as if $\widehat{\mathcal{C}}$ was a real variable.

One would not expect the result $\psi_{\rm phys}$ to be a normalizable function (simply because in general there are no normalizable solutions to the wave equation), but one may expect it to be a function on $\mathcal{M}$ solving (2) — at least for reasonable choices of ψ.

Since an element $\psi \in \mathcal{H}$ does not satisfy (2) and hence does not represent a physical state of the system, it is called "kinematical state" and $\mathcal{H}$ is called "kinematical Hilbert space". Accepting the way (8) of writing down the relation between kinematical and physical states, we proceed and try to compute the inner product between two physical states,

$$\psi_{1,\rm phys} = \delta(\widehat{\mathcal{C}})\,\psi_1 \qquad\qquad \psi_{2,\rm phys} = \delta(\widehat{\mathcal{C}})\,\psi_2\,, \tag{9}$$

a quantity about which we know from the outset that it is ill-defined: Formally, we get

$$\langle\psi_{1,\rm phys}, \psi_{2,\rm phys}\rangle = \langle\psi_1, \delta(\widehat{\mathcal{C}})^2\psi_2\rangle = \delta(0)\,\langle\psi_1\delta(\widehat{\mathcal{C}})\psi_2\rangle \tag{10}$$

where I have written $\delta(\widehat{\mathcal{C}})^2 = \delta(0)\delta(\widehat{\mathcal{C}})$ in order to indicate where the infinity comes from.

The key idea with far reaching consequences is now very simple: just drop $\delta(0)$! Given the two states (9), define an inner product $\langle\,,\,\rangle_{\rm phys}$ as

$$\langle\psi_{1,\rm phys}, \psi_{2,\rm phys}\rangle_{\rm phys} = \langle\psi_1, \delta(\widehat{\mathcal{C}})\,\psi_2\rangle\,. \tag{11}$$

This has a chance to be finite even if the two physical states are not normalizable with respect to the kinetical inner product $\langle\,,\,\rangle$.

The equations written down so far provide the motivation (and the basis for the exactification) of the RAQ scheme in the case of a single constraint. The set of solutions to (2) which may be written in the form (8) and which are normalizable with respect to the inner product (11) defines the Hilbert space $(\mathcal{H}_{\rm phys}, \langle\,,\,\rangle_{\rm phys})$ of physical states.

There is another way to write down the same construction. The quantum constraint (2) may be viewed as the eigenvalue problem of the constraint operator $\widehat{\mathcal{C}}$ for the eigenvalue zero. If zero was indeed an eigenvalue of $\widehat{\mathcal{C}}$, then a solution $\psi_{\rm phys}$ would just be an element of the kernel of $\widehat{\mathcal{C}}$. Then one could interpret (8) as the

projection onto this kernel (and $\delta(\widehat{\mathcal{C}})$ would in fact better be written as $\delta_{\widehat{\mathcal{C}},0}$). However, since $\widehat{\mathcal{C}}$ has purely continuous spectrum near zero, there is no such element in the Hilbert space $\mathcal{H}$, and zero is only a *generalized* eigenvalue. One thus needs a generalized sort of "projection" onto the (generalized) kernel of $\widehat{\mathcal{C}}$. This throws one *out of* the kinematical Hilbert space $\mathcal{H}$. In the approach outlined above the quantity $\delta(\widehat{\mathcal{C}})$ has played the role of such an operation. Again, some acrobatics with the delta function provides an intuitive piece of insight. Under quite weak conditions there is a set of generalized eigenvectors $\psi_{j,E}$ of $\widehat{\mathcal{C}}$, where E is a continuous and j a discrete index, satisfying

$$\begin{aligned}\widehat{\mathcal{C}}\,\psi_{j,E} &= E\,\psi_{j,E} &(12)\\ \langle\psi_{j,E},\psi_{j',E'}\rangle &= \delta(E-E')\,\delta_{jj'}\,, &(13)\end{aligned}$$

and being complete in the sense that any $\psi\in\mathcal{H}$ is given by a linear superposition of the form

$$\psi = \sum_j \int dE\, c_j(E)\,\psi_{j,E} \tag{14}$$

(I omit measure-theoretical subtleties here). The discrete index accounts for the degeneracy of the generalized eigenvalues E. (One may think of the Fourier transformation in $L^2(\mathbb{R})$ for a well-known example of a basis of generalized eigenvectors).

Note that the $\psi_{j,E}$ are not elements of $\mathcal{H}$. Nevertheless, by virtue of (12), the functions $\psi_{j,0}$ solve the quantum constraint (2). (This becomes quite clear if one likes to compute the $\psi_{j,E}$ in practice: one would have to *solve* (13) for any E (up to boundary conditions that essentially exclude exponentially increasing functions), the index j just counting the linearly independent solutions. In this sense, the physical states (i.e. the solutions of the wave equation) are already at our disposal. The formal inner product (13) between two $E=0$ functions of course diverges

$$\langle\psi_{j,0},\psi_{j',0}\rangle = \delta(0)\ \delta_{jj'}\,, \tag{15}$$

and this can be cured by dropping $\delta(0)$ and *defining* the physical inner product as

$$\langle\psi_{j,0},\psi_{j',0}\rangle_{\rm phys} = \delta_{jj'}. \tag{16}$$

The $\delta(0)$ from (10) and from (15) are in fact the same! Again, in this picture, the non-normalizability of physical states with respect to the kinetical inner product has been overcome by dividing by an infinite constant. The set $\{\psi_{j,0}\}$ then provides an orthonormal basis (in the usual sense) of $\mathcal{H}_{\rm phys}$. In order to make this second approach more rigorous one has to make use of the spectral theory of self-adjoint linear operators. In fact, this one of Marolf's contributions to the RAQ scheme [6]). The relation to the first approach is provided by $\psi_{j,E}=\delta(\widehat{\mathcal{C}}-E)\chi_{j,E}$ for suitably chosen kinematical states $\chi_{j,E}$. For $E=0$, this reduces to (8).

The framework outlined so far in this Section, although at a heuristic level, suffices to carry out practical computations. For non-trivial cases, the wave equation is of course very difficult to solve explicitly, and so a closed formula for the physical inner product is still out of reach for many (actually most) interesting models. But these difficulties are more of practical than of conceptual nature.

An instructive example is that of the flat Klein-Gordon equation ($g_{\alpha\beta} = \eta_{\alpha\beta}$ flat and $U = m^2$ constant, as treated in Section 1). Solutions of the wave equation are functions of the form (4), and a more or less straightforward application of the formulae written down so far (either by computing the action of $\delta(\widehat{\mathcal{C}})$ or by finding $\psi_{j,E}$) reveals that the physical inner product between two such functions ψ_1, ψ_2 is given by

$$\langle \psi_1, \psi_2 \rangle_{\rm phys} = 4\pi \int d^3k \Big(a_1^*(\vec{k}) a_2(\vec{k}) + b_1^*(\vec{k}) b_2(\vec{k}) \Big) \,. \tag{17}$$

This provides the exact definition of the physical Hilbert space $\mathcal{H}_{\rm phys}$: it consists of those functions of the form (4) which are normalizable with respect to (17). Both $a(\vec{k})$ and $b(\vec{k})$ must be square integrable, so that $\mathcal{H}_{\rm phys}$ becomes isomorphic to $L^2(\mathbb{R}) \times L^2(\mathbb{R})$. Note that the difference between (5) and (17) is essentially a sign. The physical inner product is thus obtained by reversing the sign of the indefinite scalar product (3) in the negative frequency sector. Although this may be disturbing, such a reversion of sign is perfectly compatible with Lorentz invariance. (In fact, this observation was one of the starting points of Higuchi's contribution to the RAQ scheme [3]). The positivity of the inner product $\langle \, , \, \rangle_{\rm phys}$, being desirable from the point of view of interpretation, provides another major advantage over the indefinite scalar product Q: it leads to a precise definition of the set of states to work with, whereas Q does not (as already remarked in Section 1: the set of solutions ψ to the wave equation for which $Q(\psi, \psi)$ is finite contains asymptotically ill-behaved functions).

There is yet a third way to intuitively understand the transition from the kinetical to the physical inner product. Classically, the constraint function (1) plays a double role, the first of which — constraining the phase space variables and thus leading to the reduced phase space — is carried over to the quantum theory by imposing the wave equation (2). However, its second role is to generate gauge transformations (as any first class constraint does), the infinitesimal expression for which is given by $\delta f = \epsilon\{\mathcal{C}, f\}$ for functions $f \equiv f(x,p)$. This induces a flow on the reduced phase space, connecting points which have to be considered physically equivalent. Quantum mechanically, this is accounted for by inserting a gauge-fixing delta function together with a Faddeev-Popov determinant into the integral (6). The kinetical inner product — being defined without this insertion — contains a redundant piece, namely the integration over gauge-orbits. It thus differs from the physical inner product by a factor which represents the redundant integration. The latter is infinite and may in turn be identified with our $\delta(0)$. It was stressed by Woodard [7] that this procedure — being standard reasoning in perturbation theory around flat space — should be

applied in the quantum gravity and quantum cosmology context as well[1].

A way of interpreting the physical states constructed by the RAQ scheme in terms of observations has been proposed by Marolf [8]. I will not enter this issue, nor will I say something about how classical observables are promoted into quantum ones, i.e. self-adjoint linear operators on $\mathcal{H}_{\rm phys}$ (see Ref. [5] for this topic), but I will now extract a more general structure from the picture developed so far. By separating several issues from each other it should also become clear how the scheme can be extended to more realistic situations.

3 Towards the general RAQ scheme — still hand-waving

The RAQ scheme is in fact a list of rules that may sometimes be applied more or less straightforwardly (as for the simple models considered above, given that its ingredient are not too pathological), while in more sophisted systems crucial ambiguities may appear, so that a lot of choice has to be made on the way. In order to outline two major items of this list, I begin reformulating what we have done above in a slightly more precise language. (I ignore here many problems that arise even *before* this point, starting from a quite generic classical system, for example the very definition of the kinematical Hilbert space in cases where the analogy of (6) is not obvious; see Refs. [1][5]).

The minisuperspace framework provides the kinematical Hilbert space $\mathcal{H}$ and the constraint operator $\widehat{\mathcal{C}}$ as objects to start with. In more sophisticated models one will encounter a kinematical Hilbert space $\mathcal{H}$ as well, together with a family of constraint operators (say $\widehat{\mathcal{C}}_I$). The goal is to "solve the constraint(s)" by defining a Hilbert space of physical states *just in terms of the Hilbert space framework*. So let us reconsider the minisuperspace example, trying to extract a procedure that can be stated solely in terms of $(\mathcal{H}, \widehat{\mathcal{C}})$ and structures contained in this framework.

First of all we note that the constraint operator $\widehat{\mathcal{C}}$ is an unbounded linear operator, so it can be defined only on a dense subspace $\mathcal{D}(\widehat{\mathcal{C}})$ of $\mathcal{H}$,

$$\widehat{\mathcal{C}} : \mathcal{D}(\widehat{\mathcal{C}}) \to \mathcal{H} \,. \tag{18}$$

Nevertheless, if it is self-adjoined, the operators $e^{i\lambda\widehat{\mathcal{C}}}$ are unitary and hence well-defined on the whole of $\mathcal{H}$.

Almost everything said in the preceding Section cries for an interpretation of the physical states in terms of distributions. The object $\delta(\widehat{\mathcal{C}})\,\psi$ — which I have argued to be a physical state, i.e. a function $\psi_{\rm phys}$ on $\mathcal{M}$ satisfying the quantum

[1] I became aware of Woodard's beautiful paper only after having written up this conference contribution, thanks to a hint by Steve Carlip.

constraint (2) — is not normalizable with respect to $\langle\, ,\, \rangle$, hence it is not an element of the kinematical Hilbert space $\mathcal{H}$. However, the quantity $\langle \delta(\widehat{\mathcal{C}})\, \psi, f \rangle$, when witten as an integral, using (7), has a chance to be finite for at least a large set of elements $f \in \mathcal{H}$ (just as the integral $\int dx\, f(x)$ has a chance to be finite for certain elements $f \in L^2(\mathbb{R})$). In general, the introduction of distributions on a Hilbert space may be motivated by the attempt to give the scalar product between a non-normalizable object (distribution) and a normalizable object (test function) a precise meaning. In our case, given ψ, we encounter the linear assignment

$$\eta(\psi) : f \;\mapsto\; \langle \delta(\widehat{\mathcal{C}})\, \psi, f \rangle = \int_\infty^\infty \frac{d\lambda}{2\pi}\, \langle \psi, e^{-i\lambda\widehat{\mathcal{C}}} f \rangle\,. \tag{19}$$

Here, ψ stands for a kinematical state and $\eta(\psi)$ is a linear map from (a subspace of) $\mathcal{H}$ into the complex numbers. *This map* encodes all the information I have naively written down as $\delta(\widehat{\mathcal{C}})\, \psi$, so that the idea is to *consider it as a physical state*. If this can be made rigorous, η would be an anti-linear map sending (certain) kinematical states to physical states.

This means that we re-interpret the *functions* $\psi_{\rm phys}$ satisfying (2) as *maps* $\eta(\psi)$ of the type (19). In this way physical states become distributions. These distributions may of course still be *represented as functions* $\psi_{\rm phys}$ (the relation between the function $\psi_{\rm phys}$ and the distribution $\eta(\psi)$ being equation (19), when written in the form $\eta(\psi) : f \mapsto \langle \psi_{\rm phys}, f \rangle$, which is just an integral over $\mathcal{M}$ of the form (6)), so that the arguments of the preceding Section are not invalidated. However, the new point of view is generalizable since it develops the notion of non-normalizable objects *from the Hilbert space itself* and does not need to refer to specific properties of the model.

What spaces are ψ and f allowed to come from? After all, the expression (19) shall be finite. If a large subspace of $\mathcal{H}$ is allowed to inhabit the possible ψ's, there will be only few f's for which (19) exists. Conversely, if ψ is restricted to come from a small subspace, then we will find many f's for which (19) exists. Here things become ambigous, although any reasonable choice will yield the same final result for our simple minisuperspace models. However, instead of searching a preferred choice of these spaces just for the models considered so far, we anticipate more sophisticated situations and try to extract the crucial contents of the present situation. Let us denote by Φ the subspace of $\mathcal{H}$ in which ψ has to lie, and suppose it to be dense in $\mathcal{H}$.

The *first* major choice in the RAQ program is thus to fix a dense subspace Φ of $\mathcal{H}$. Moreover, the space Φ shall be a subspace of $\mathcal{D}(\widehat{\mathcal{C}})$ and it shall be invariant under the action of the constraint operator $\widehat{\mathcal{C}}$. Let Φ' be the (topological) dual of Φ, i.e. the space of all continuous linear map from Φ into the complex numbers. This is the space of distributions which provides a home for the generalized eigenvectors of $\widehat{\mathcal{C}}$ and inside which we will identify the physical states. In order to have a non-trivial notion of continuity, we must assume a topology on Φ which is finer than

the Hilbert-space topology. It is usually assumed to be a nuclear topology. For simple models like (2), one does not really need to postulate so much freedom, but in complicated models this might be necessary (although the last word has certainly not yet been spoken on this issue). Anyway, given such a suspace, one may imbed $\mathcal{H}$ in Φ' (because taking the inner product with a fixed element of $\mathcal{H}$ defines an element of Φ'), so that one gets the chain of inclusions $\Phi \subset \mathcal{H} \subset \Phi'$ (which is called a Gelfand triple).

The elements of Φ will play the role of test functions. Possible candidates for Φ to start with in minisuperspace models are thus of the Schwarz or C_0^∞ type, whereas in complicated situations, e.g. in all attempts to treat full gravity, the choice of Φ can be a difficult task.

The *second* major choice to be made in the RAQ program is to fix an anti-linear map (called "rigging map")

$$\eta : \Phi \to \Phi' \tag{20}$$

such that $\eta(\psi)$ "satisfies the constraint" for any $\psi \in \Phi$. What does this mean? In the simple models considered above, the constraint is solved because the quantity $\delta(\widehat{\mathcal{C}})$ is floating around. Equation (19) is just the definition of η, once the question of domains has been fixed. It may be interpreted by stating that the constraint operator $\widehat{\mathcal{C}}$ generates a group (of unitary operators) over which an average is performed. (Hence the alternative name "group average" for RAQ. There should, by the way, also be relations to the Faddeev-Popov way to define the physical inner product mentioned in Section 2). This provides a hint how to proceed in case of more than just one constraint: Consider the group (of unitary operators) generated by the constraints and carry out the average. The average over a compact group is provided by the (unique) Haar measure, in which case (19) finds a natural generalization. However, in case a non-compact group appears, there is no general prescription, and the problem of averaging must be tackled for any particular model. (For example, since the diffeomorphism group on a manifold is not compact, the diffeomorphism constraint in loop quantum gravity is of this type [1][9]).

In order not to refer to an average procedure that might not exist in certain models, I rather state the *goal* which has to be achieved and for which the group average is just a tool: solving the constraint(s). Since we are now considering a physical state to be an element $\eta(\psi) \in \Phi'$ (stemming from an element $\psi \in \Phi$), the statement that it solves the constraint is reformulated as

$$(\eta(\psi))\, \widehat{\mathcal{C}} f = 0 \qquad\qquad \forall\, f \in \Phi\,. \tag{21}$$

Note that, since Φ is left invariant by $\widehat{\mathcal{C}}$, this equation makes perfect sense. Thus, expressed in Dirac's notation, one does not actually solve the equation $\widehat{\mathcal{C}}\,|\psi\rangle = 0$ but rather the equation $\langle\psi|\,\widehat{\mathcal{C}} = 0$ in a distributional sense. (This reflects, by the way, the anti-linear nature of η). In the minisuperspace models this condition is identical to $\psi_{\rm phys}$ satisfying the wave equation (just partially integrate $\langle\psi_{\rm phys}, \widehat{\mathcal{C}} f\rangle = 0\,\forall f \in \Phi$),

so that the hand-waving framework of the preceding Section is recovered. Also, the generalization of (21) to the case of several constraints is obvious, once their action is known. (All constraints shall act on Φ and leave Φ invariant). In general, whenever a group average exists, the resulting map η satisfies this property. If a group average is not obviously defined, one might try to achieve (21) for all $\psi \in \Phi$ by any other method at hand. In order to mention an example: for the diffeomorphism constraint in loop quantum gravity (which is actually an uncountable family of constraints), it is an easier task to realize an appropriate version of (21) than to try to generalize the notion of group average.

There are some further requirements for η, among which the most important ones are $(\eta(\psi_1))\psi_2 = ((\eta(\psi_2))\psi_1)^*$ and $(\eta(\psi))\psi \geq 0$. All this may easily be satisfied for the models of the type (2) but prevents potential difficulties for more complicated scenarios.

Once having succeeded to fix Φ and η, only some simple definitions remain to be done. The physical Hilbert space is defined as the completion of $\eta(\Phi)$ with respect to the inner product

$$\langle \eta(\psi_1), \eta(\psi_2) \rangle_{\rm phys} = (\eta(\psi_2))\,\psi_1 \tag{22}$$

for $\psi_1, \psi_2 \in \Phi$. Note that ψ_1 and ψ_2 have interchanged position at the right hand side. This stems from the anti-linearity of η, together with the convention that $\langle\, ,\, \rangle_{\rm phys}$ shall be anti-linear in the first factor. The fact that there is just one η at the right hand side is the exactified version of dropping an infinite factor $\delta(0)$ as was done in Section 2. In this way the space of physical states emerges on a rather abstract level as a certain set of distributions solving the constraint(s) in a certain sense. In the minisuperspace models, as already anticipated, the elements of $\eta(\Phi)$ may be represented as functions $\psi_{\rm phys}$ on $\mathcal{M}$ satisfying the wave equation (at least for a reasonable choice of Φ), whereas passing over to the completion adds objects which satisfy (2) only in a Hilbert space sense (thus implying a weaker notion of differentiability). In more complicated models, such an intuitive procedure might not be possible, in which case one will be obliged to work entirely at the abstract level in order to arrive at $\mathcal{H}_{\rm phys}$.

This is not the end of the story. I have ignored the problem of observables and many technicalities. In a realistic system like full gravity, any step in the list (including those I did not even mention) provides its own problems. Despite the individuality and complexity of these problems, there are some common underlying ideas, encoded in the RAQ scheme, some of which I (hopefully) have made clear.

References

[1] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão and T. Thiemann, "Quantization of diffeomorphism invariant theories of connections with local degrees of freedom", *J. Math. Phys.* **36**, 6456 (1995).

[2] O. Nachtmann, "Dynamische Stabilität im de-Sitter-Raum", *Sitz. Ber. Öst. Ak. d. Wiss. II* **176**, 363 (1968); "Continuous Creation in a Closed World Model", *Z. Physik* **208**, 113 (1968).

[3] A. Higuchi, "Quantum linearization instabilities of de Sitter spacetime: II", *Class. Quantum Grav.* **8**, 1983 (1991); "Linearized quantum gravity in flat space with toroidal topology", *Class. Quantum Grav.* **8**, 2023 (1991); "Quantum Linearization Instabilities of de Sitter Spacetime", *in*: B. L. Hu and T. A. Jacobson (eds.), *Directions in General Relativity, Vol. 2*, Cambridge University Press (Cambridge, 1993), p. 146.

[4] N. P. Landsman, "Rieffel induction as generalized quantum Marsden-Weinstein reduction", *J. Geom. Phys.* **15**, 285 (1995); N. P. Landsman and U. A. Wiedemann, "Massless particles, electromagnetism, and Rieffel induction", *Rev. Mod. Phys.* **7**, 923 (1995).

[5] "Refined algebraic quantization: Systems with a single constraint", *preprint* gr-qc/9508015, to appear in *Banach Center Publications.*

[6] D. Marolf, "The spectral analysis inner product for quantum gravity", *preprint* gr-qc/9409036, to appear in the Proceedings of the VIIth Marcel-Grossman Conference, R. Ruffini and M. Keiser (eds) (World Scientific, Singapore, 1995).

[7] R. P. Woodard, "Enforcing the Wheeler-DeWitt constraint the easy way", *Class. Quantum Grav.* **10**, 438 (1993).

[8] D. Marolf, "Quantum Observables and Recollapsing Dynamics", *Class. Quantum Grav.* **12**, 1199 (1995).

[9] D. Marolf, J. Mourão and T. Thiemann, "The Status of Diffeomorphism Superselection in Euclidean 2+1 Gravity", *preprint* gr-qc/9701068, to appear in *J. Math. Phys.*